



**University of  
Leicester**

**DEPARTMENT OF ECONOMICS**

# **The Nonexistence of Instrumental Variables**

**Stephen Hall, University of Leicester, UK**

**George S. Tavlas, Bank of Greece, Economics Research Department, Greece**

**P. A. V. B. Swamy, Federal Reserve Board (Retired), USA**

**Working Paper No. 11/24**

**March 2011**

# **The Nonexistence of Instrumental Variables**

Stephen G. Hall <sup>c,b</sup>, George S Tavlás <sup>b</sup>, and P.A.V.B. Swamy <sup>a</sup>

<sup>a</sup> *Board of Governors of the Federal Reserve System (retired), Washington, DC, USA*

<sup>b</sup> *Bank of Greece, 21, El, Venizelos Ave, 102 50 Athens, Greece. Tel.++30210 320 2370; Fax: ++30210 320 2432. Email: [gtavlas@bankofgreece.gr](mailto:gtavlas@bankofgreece.gr)*

<sup>c</sup> *Dept of Economics, University of Leicester*

## **Abstract**

The method of instrumental variables (IV) and the generalized method of moments (GMM) has become a central technique in health economics as a method to help to disentangle the complex question of causality. However the application of these techniques require data on a sufficient number of instrumental variables which are both independent and relevant. We argue that in general such instruments cannot exist. This is a reason for the widespread finding of weak instruments..

*JEL classification:* C32, C51

## **1. Introduction**

Researchers are becoming increasingly aware that there are often serious problems with the use of instrumental variable based techniques (both instrumental variable (IV) estimation and versions of generalized methods of moments (GMM) which use instrumental variables). Auld(2006) points out that while many researchers in health economics resort to instrumental variable techniques to identify causal effects there are serious difficulties with finding valid instruments in many contexts. A valid instrument must be uncorrelated with the errors in an equation (independent) and correlated with the variable which it is being used to replace (relevant). Increasingly, researchers are finding that when a set of instruments are independent of

the error they often have little relevance, this is the well known problem of ‘weak instruments’. Recently Basu, Heckman, Navarro-Lozano and Urzua(2007), following on from earlier work by Imbens and Angrist (1994) have shown that there can also be serious problems with the use of instrumental variables where treatment effects are heterogeneous across patients. In this case researchers aim to achieve an average treatment effect (ATE) but IV does not provide a consistent estimate of this and instead only estimates a local average treatment effect (LATE), see Jones’s (2009) survey for many examples of IV applied in the Health Economics context and for a thorough discussion of ways to estimate a causal relationship. This argument is similar, although less general, to the criticism of IV made below. Here we show that any type of model misspecification may be represented by a model without errors but with coefficients which vary across each individual or each time period (we refer to time here but by simply replacing  $t$  with  $i$  the analysis goes through exactly with respect to cross section data and with only a minor complication panels may also be included in the proofs below). Once this has been done then the problems identified by Basu et al become relevant to all uses of IV.

We show that, in general, valid instrumental variables cannot, in principle exist. We do this by first developing a general representation of how misspecification arises in an econometric model, in section 2, and then showing that this representation implies that valid instrumental variables cannot exist. In section 3 we then illustrate this argument in the context of a very simple model involving measurement error. Section 4 concludes.

## **2. A General Representation of Misspecification**

Economic theory suggests relationships between variables but it does not usually give clear guidance as to the correct functional form or the exact set of variables which might be relevant.

For example, consider an economic variable, denoted by  $y_t^*$ , and its determinants, denoted by  $x_{jt}^*, j = 1, \dots, L_t$ . Here the total number  $L_t$  of determinants may be time dependent. Typically, data on  $y_t^*$  and on a subset  $K - 1$  of the  $L_t$  determinants are available. The remaining  $L_t - K + 1$  determinants are omitted from the model either because they are unobserved or for some other reason. Moreover, these data may contain measurement errors. Let  $y_t = y_t^* + v_{0t}$  and  $x_{jt} = x_{jt}^* + v_{jt}, j = 1, \dots, K - 1$ , where the variables without an asterisk are observable, the variables with an asterisk are unobservable, and  $v$  s are measurement errors. The theoretical relationship is

$$y_t^* = f_t(x_{1t}^*, \dots, x_{L_t}^*) \quad (t = 1, \dots, T) \quad (1)$$

with unknown functional form.

Without misspecifying the relationship in (1), we can write

$$y_t^* = \alpha_{0t} + \sum_{j=1}^{K-1} \alpha_{jt} x_{jt}^* + \sum_{g=K}^{L_t} \alpha_{gt} x_{gt}^* \quad (2)$$

where the time profiles of the coefficients are determined by the correct functional form of model (1). These time profiles are unknown, since the correct functional form is unknown<sup>1</sup>. Allowing the coefficients of equation (2) to vary freely defines an infinite class of functional forms, which surely encompasses the correct (but unknown) functional form of (1) as a special case.

The correlations between the omitted determinants and the observed determinants are implied by

$$x_{gt}^* = \lambda_{0gt} + \sum_{j=1}^{K-1} \lambda_{jgt} x_{jt}^* \quad (g = K, \dots, L_t) \quad (3)$$

where  $\lambda_{0gt}$  is a portion of  $x_{gt}^*$  remaining after the effects of the  $x_{jt}^*$  's have been removed from

---

<sup>1</sup> It is possible to represent any functional form exactly by a time varying Parameter model, we refer to this as the Swamy theorem. Recently, Granger(2008) has also confirmed this theorem although he attributes the proof to Halbert White.

$x_{gt}^*$ . Since we do not have data on the  $L_t - K + 1$   $x_{gt}^*$  variables, we can eliminate them from equation (2) by substituting (3) into (2), which gives

$$y_t^* = \alpha_{0t} + \sum_{g=K}^{L_t} \alpha_{gt} \lambda_{0gt} + \sum_{j=1}^{K-1} (\alpha_{jt} + \sum_{g=K}^{L_t} \alpha_{gt} \lambda_{jgt}) x_{jt}^* \quad (4)$$

Note that equation (4) shows  $y_t^*$  as a function of  $K - 1$  included determinants and the remainders of the excluded variables - - i.e., what remains after subtracting the effects on the excluded variables of the  $K - 1$  observable determinants. Equation (4) accounts for both the unknown functional form (since it is derived from equation (2)) and the full set of (time-varying) determinants of  $y_t^*$ . It does not, however, account for measurement errors. In this connection, consider model (2) again. It is not in an estimable form. Such a form is derived below.

In terms of the observable variables, equation (2) can be written as

$$y_t = \gamma_{0t} + \sum_{j=1}^{K-1} \gamma_{jt} x_{jt} \quad (5)$$

We call the  $x_{gt}^*$ 's "excluded variables" because they are excluded from model (5). The  $x_{jt}$ 's are the included explanatory variables. Model (5) coincides with model (2) if

$$\gamma_{0t} = \alpha_{0t} + \sum_{g=K}^{L_t} \alpha_{gt} \lambda_{0gt} + v_{0t} \quad (6)$$

$$\gamma_{jt} = (\alpha_{jt} + \sum_{g=K}^{L_t} \alpha_{gt} \lambda_{jgt}) (1 - \frac{v_{jt}}{x_{jt}}) \quad (j = 1, \dots, K-1) \quad (7)$$

These equations are derived by establishing the correspondence between equations (4) and (5).<sup>2</sup>

The terms on the right-hand side of equations (6) and (7) provide crucial information. Equation (4) shows that the  $\lambda_{0gt}$ 's in conjunction with the  $x_{jt}^*$ 's are at least sufficient to

---

<sup>2</sup> For the derivation, see Swamy and Tavlás (2007).

determine  $y_t^*$ . The interpretation of the terms on the right-hand side of equation (7) and their implications are as follows:

- The term  $\alpha_{jt}$  is equal to  $\partial y_t^* / \partial x_{jt}^*$  (if  $y_t^*$  is a continuous function of  $x_{jt}^*$ ) and corresponds to the bias-free effect of  $x_{jt}^*$  on  $y_t^*$ , as can be seen from (2).
- The term  $\sum_{g=K}^{L_t} \alpha_{gt} \lambda_{jgt}$  measures omitted-variables bias. Note that each term in this sum is the products of two coefficients - - the effect of the excluded variable  $x_{gt}^*$  on  $y_t^*$  (i.e.,  $\alpha_{gt}$ ) and the effect of the included variable  $x_{jt}^*$  on the excluded variable  $x_{gt}^*$  (i.e.,  $\lambda_{jgt}$ ).
- The term  $(\alpha_{jt} + \sum_{g=K}^{L_t} \alpha_{gt} \lambda_{jgt})(-v_{jt} / x_{jt})$  measures measurement-errors bias.<sup>3</sup> These biases exist whenever estimates of some theoretical variables are used as explanatory variables.
- The explanatory variables of model (5) are correlated with their own coefficients because the measurement-error bias component of  $\gamma_{jt}$  is a function of  $x_{jt}$ .

Having derived the model in (5) which explicitly includes all these forms of misspecification it is now possible to show why valid instruments can not be found for this model. Under IV or GMM we are imposing constant parameters on (5) we can therefore re-write 5 as;

$$y_t = \alpha_{0t} + \sum_{g=K}^{L_t} \alpha_{gt} + v_{0t} + \sum_{j=1}^{K-1} (\alpha_{jt} + \sum_{g=K}^{L_t} \alpha_{gt} \lambda_{jgt}) (1 - \frac{v_{jt}}{x_{jt}}) x_{jt} \quad (8)$$

And then imposing a constant parameter model we get

---

<sup>3</sup> The minus sign in the expression reflects the fact that the second parenthetical term on the right-hand side of (7) in one minus the ratio  $(v_{jt} / x_{jt})$ .

$$y_t = \beta_0 + \sum_{j=1}^{K-1} \beta_j x_{jt} + (\alpha_{0t} + \sum_{g=K}^L \alpha_{gt} + v_{0t} - \beta_0) + \sum_{j=1}^{K-1} ((\alpha_{jt} + \sum_{g=K}^{L_t} \alpha_{gt} \lambda_{jgt}) (1 - \frac{v_{jt}}{x_{jt}}) - \beta_j) x_{jt} \quad (9)$$

Where the last two terms in (9) become the error term in the model. The problem with instrumental variables in this context now becomes apparent; we need to find a variable which is correlated with  $x_{jt}$  but which is not correlated with the error term, which itself contains  $x_{jt}$ . Such a variable cannot in general exist. The only case in which a valid instrument may exist is when  $\lambda_{jgt} = 0$  For all  $j, g$  and  $t$ , which implies that there is no correlation between the included and excluded variables at any point in time, or that there are no omitted variables. And that  $\alpha_{jt} = \beta_j$  for all  $t$ , which implies that there is no time variation in the true parameter which, in turn, implies that we have the correct functional form. This amounts to saying that there is no misspecification in the model and, hence, there is no need for instruments in the first place.

### 3. A Simple Example

Consider a very trivial example where the only misspecification is measurement error in the independent variable. Assume we have a perfectly fitting relationship in the true variables

$$Y^* = \beta X^* \quad (10)$$

where the measured value of  $X$  is given by

$$X = X^* + v \quad (11)$$

Then the model we estimate is

$$Y^* = \beta^1 X + \varepsilon \quad (12)$$

where  $v$  and  $\varepsilon$  are error terms. Then there are two ways we can demonstrate the problem with IV applied to (12). First we may consider the issue from a TVC perspective and we may write an exact version of 10 as

$$Y^* = \beta_t X \quad (13)$$

Then if we apply a fixed parameter model to this equation we get

$$Y^* = \beta^2 X + (\beta_1 - \beta^2) X \quad (14)$$

And the last term is the error term in (12) and we can see that no valid instruments can exist for  $X$  since  $X$  is also in the error term. We can also see the same problem from a more conventional perspective. If we perform a fixed parameter regression then we can rewrite (12) as

$$Y^* = \beta^1 X + (\beta X - \beta^1 X) \quad (15)$$

Where again the term in brackets in (15) is the error term in (12) and we can see that the error term again contains the same variable that we are trying to instrument and so no valid instrument can exist.

#### 4. Conclusion

The instrumental variables that are correlated with the  $x_{jt}$ 's of model (5) but not with the error terms of model (9) do not in general exist because these error terms also involve the  $x_{jt}$ 's. These arguments make it clear why practical work with IV methods is plagued by the problem of weak instruments. The general difficulty with IV has become increasingly obvious in Health Economics as well as other areas of our discipline Basu et al (2007). We have argued here that the problems of IV are endemic to all models not just to certain ones such as those with heterogeneous coefficients. We would argue that a much better way forward in terms of practical estimation rests on a recognition of all the potential sources of misspecification which are apparent in (5) and which starts from a time varying coefficient model as outlined in Swamy, Hall, Tavlas and Hondroyianis(2010).

#### Acknowledgments



The views expressed in this paper are the authors' own and are not to be interpreted as those of their respective past or present institutions.

## References

Auld C. M. ( 2006 ) Using Observational Data to Identify the Causal effects of Health-related behavior. In Jones A.M. (ed), *The Elgar Companion to Health Economics*, Aldershot: Edward Elgar

Basu A., J.J. Heckman, S. Navarro-Lozano and S. Urzua. (2007) Use of Instrumental Variables in the Presence of Heterogeneity and Self-Selection: An Application To Treatments of Breast Cancer Patients. *Health Economics*, 16: 1133-1157

Granger,(2008) C.W.J.'Non-Linear Models: Where do we go next – Time Varying Parameter Models? *Studies in Nonlinear Dynamics and Econometrics*, 12, 3, 1-9.

Greene, W.H., 2008, *Econometric analysis*, 6<sup>th</sup> edn. (Pearson/Prentice Hall, Upper Saddle River, New Jersey).

Hall, S.G., G. Hondroyiannis, P.A.V.B. Swamy and G.S. Tavlas, 2009, The new Keynesian Phillips curve and lagged inflation: A case of spurious correlation? *Southern Economic Journal*.

Hondroyiannis, G., P.A.V.B. Swamy, and G.S. Tavlas, 2009, The new Keynesian Phillips curve in a time-varying coefficient environment: Some European evidence, *Macroeconomic Dynamics* 13, 149-166.

Imbens G, J. Angrist (1994) Identification and Estimation of Local Average Treatment Effects, *Econometrica* 62: 467-475.

Jones A.M., (2009) Panel Data Methods and Applications to Health Economics, in the *Palgrave Handbook of Econometrics*, vol 2 (eds) T.C. Mills and K.D. Patterson, Palgrave Macmillan, London.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl and T. Lee, 1985, *The theory and practice of econometrics*, 2<sup>nd</sup> edn. (John Wiley and Sons, New York).

Pratt, J.W. and R. Schlaifer, 1988, On the interpretation and observation of laws, *Journal of Econometrics* 39, 23-52.

Swamy P.A.V.B and G.S.Tavlas, 2001, Random Coefficient Models, Ch 19. In Baltagi B.H.(ed.) *A Companion to Theoretical Econometrics*, Malden, Blackwell.

Swamy, P.A.V.B. and G.S. Tavlas, 2007, The new Keynesian Phillips curve and inflation expectations: Re-specification and interpretations, *Economic Theory* 31, 293-306.

Swamy,P.A.V.B.,S.G. Hall, G.Tavlas and G.Hondroyiannis (1020) "Estimation of Parameters in the Presence of Model Misspecification and Measurement Error," Studies in Nonlinear Dynamics & Econometrics.